

STRONG CONVERGENCE OF A FULLY DISCRETE FINITE ELEMENT APPROXIMATION OF THE STOCHASTIC CAHN–HILLIARD EQUATION

DAISUKE FURIHATA, MIHÁLY KOVÁCS, STIG LARSSON, AND FREDRIK LINDGREN

ABSTRACT. We consider the stochastic Cahn–Hilliard equation driven by additive Gaussian noise in a convex domain with polygonal boundary in dimension $d \leq 3$. We discretize the equation using a standard finite element method in space and a fully implicit backward Euler method in time. By proving optimal error estimates on subsets of the probability space with arbitrarily large probability and uniform-in-time moment bounds we show that the numerical solution converges strongly to the solution as the discretization parameters tend to zero.

1. INTRODUCTION

Let $\mathcal{D} \subset \mathbb{R}^d$, $d \leq 3$, be a convex spatial domain with polygonal boundary $\partial\mathcal{D}$ and consider the stochastic Cahn–Hilliard equation, also known as the Cahn–Hilliard–Cook equation [4, 8, 10], written in the abstract Itô form

$$(1.1) \quad dX + A(AX + f(X)) dt = dW, \quad t \in (0, T]; \quad X(0) = X_0,$$

where A is the Neumann Laplacian, $\{W(t)\}_{t \geq 0}$ is an $H := L^2(\mathcal{D})$ -valued Q -Wiener process on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t \geq 0})$ with respect to the normal filtration $\{\mathcal{F}_t\}_{t \geq 0}$. In order to avoid additional technical difficulties we assume that the initial-value X_0 is deterministic. For the solution X to preserve mass, we assume that the average $|\mathcal{D}|^{-1} \int_{\mathcal{D}} W(t) dx = 0$ for all $t \geq 0$.

The nonlinear function f is assumed to be of the form $f = F'$, where F has the following structural properties:

$$(1.2) \quad \begin{aligned} &F \text{ is a polynomial of degree 4,} \\ &F(s) \geq c_0 s^4 - c_1, \quad F''(s) \geq -\beta^2 \text{ with } c_0 > 0. \end{aligned}$$

A typical example is $F(s) = \frac{1}{4}(s^2 - \beta^2)^2$ which is a double well potential. Note that f is only locally Lipschitz and does not satisfy a linear growth condition. We also note that the restriction on the polynomial degree of F comes from the fact that we allow $d = 3$. For $d = 1, 2$ the exponent 4 in (1.2) may be replaced by any even integer larger than or equal 4 and the arguments of the paper are still valid with trivial changes. It is easy to see that properties (1.2) imply the dissipativity property

$$(1.3) \quad \langle f(x), x \rangle \geq -C_0 - C_1 \|x\|^2,$$

2000 *Mathematics Subject Classification.* 60H15, 60H35, 65C30.

Key words and phrases. stochastic partial differential equation; Cahn–Hilliard–Cook equation; additive noise; Wiener process; Finite element method, Euler method; time discretization; strong convergence.

for some $C_0, C_1 > 0$. In addition, the fact that $f''(s) \geq -\beta^2$ yields

$$(1.4) \quad F(t) - F(s) \leq f(t)(t-s) + \frac{1}{2}\beta^2(t-s)^2.$$

Finally, as f is a polynomial of degree 3 we have, for some $C > 0$, that

$$(1.5) \quad |f(x) - f(y)| \leq C(1+x^2+y^2)|x-y|.$$

It is not hard to see that, since the average of W equals 0, it follows that X preserves the average (the total mass), that is, $|\mathcal{D}|^{-1} \int_{\mathcal{D}} X(t) dx = |\mathcal{D}|^{-1} \int_{\mathcal{D}} X_0 dx$. Note that for $s_0 \in \mathbb{R}$ the function $\tilde{F}(s) = F(s+s_0)$ also satisfies the structural properties (1.2). Therefore, one can employ a change of variables $X \rightarrow X - |\mathcal{D}|^{-1} \int_{\mathcal{D}} X_0 dx$, and hence we will assume that the average $|\mathcal{D}|^{-1} \int_{\mathcal{D}} X_0 dx = 0$.

We fix a time horizon $T > 0$ and for $N \in \mathbb{N}$ consider the fully implicit finite element method

$$(1.6) \quad \begin{aligned} X_h^j - X_h^{j-1} + kA_h^2 X_h^j + kA_h P_h f(X_h^j) &= P_h \Delta W^j, \quad j = 1, 2, \dots, N, \\ X_h^0 &= P_h X_0. \end{aligned}$$

Here $k = T/N$ is the time-step, $\Delta W^j = W(t_j) - W(t_{j-1})$, A_h is the discrete Laplacian, and P_h is the orthogonal projector onto the finite element space S_h with mesh-size $h > 0$; for more details on the finite element method, see Section 2.2. It is easy to see that also X_h^j preserves the mass. The main result of the paper, Theorem 5.5, asserts that if the operator composition $A^{\frac{1}{2}}Q^{\frac{1}{2}}$ is Hilbert–Schmidt and the initial-value is regular enough; that is, for some $L > 0$,

$$|X_h^0|_1 + \mathcal{F}(X_h^0) + |A_h X_h^0 + P_h f(X_h^0)|_1 + |X_0|_1 \leq L,$$

where $\mathcal{F}(X_h^0) = \int_{\mathcal{D}} F(X_h^0) dx$ and $|v|_1 = \|A^{\frac{1}{2}}v\|$, then

$$\lim_{h,k \rightarrow 0} \mathbf{E} \sup_{t_n \in [0, T]} \|X(t_n) - X_h^n\|^2 = 0.$$

The key result used in the proof is a maximal type moment bound on $|X_h^j|_1$ which is established in Theorem 4.3 after various bootstrapping arguments. There are various difficulties in the proofs that are partly due to the finite element method. First, the finite element method is based on approximating the operator A and not A^2 . This is because the standard finite element functions belong only to the domain of $A^{\frac{1}{2}}$ but are not more regular. Loosely speaking this means that $A_h^2 \neq (A^2)_h$, which makes already the deterministic finite element analysis more challenging. Second, the presence of the finite element projection P_h in front of the semilinear term destroys some of the dissipativity properties of f . While f enjoys the dissipativity property (1.3), and even

$$\langle A^{\frac{1}{2}}f(x), A^{\frac{1}{2}}x \rangle \geq -c|x|_1^2,$$

we only have

$$\langle P_h f(v_h), v_h \rangle = \langle f(v_h), v_h \rangle \geq -C_0 - C_1 \|v_h\|^2, \quad v_h \in S_h,$$

and unfortunately

$$\langle A_h^{\frac{1}{2}} P_h f(v_h), A_h^{\frac{1}{2}} v_h \rangle = \langle A^{\frac{1}{2}} P_h f(v_h), A^{\frac{1}{2}} v_h \rangle \not\geq -c|v_h|_1^2, \quad v_h \in S_h.$$

Because of the latter we can only establish a non-uniform moment bound on $\|X_h^j\|$ in Lemma 4.2. As

$$\langle A f(x), A x \rangle \not\geq -c\|A x\|^2 \text{ and } \langle A_h P_h f(v_h), A_h v_h \rangle \not\geq -c\|A_h v_h\|^2,$$

the proof of the main moment bound in Theorem 4.3 is not based on testing (1.6) by an appropriate test function. Instead we mimic the Ljapunov functional for the original problem in the discrete setting, see (4.13). Having maximal-type moment bounds at hand we use the mild formulation of both (1.1) and (1.6) to establish pathwise error bounds on subsets of the probability space with large probability in Theorem 5.4. This turns out to be sufficient, together with some moment bounds, to show strong convergence of the numerical scheme in Theorem 5.5. Our method of proof does not give rates for the strong convergence.

Strong convergence results for numerical schemes for SPDEs with globally Lipschitz coefficients, or at least some sort of linear growth condition as in [17], are plentiful, see, for example, [9, 14, 15, 16, 19, 20, 29, 30]. In particular, in [24, 32] the linearized Cahn–Hilliard–Cook equation is treated.

In contrast, there are only few results on strong convergence of discretisation schemes for SPDEs with superlinearly growing coefficients [1, 5, 16, 18, 23, 25, 26, 31]. Furthermore, these papers dominantly establish strong convergence of various numerical schemes with no rate given (with a few exceptions), under some sort of global monotonicity assumption on the drift term which is not valid for the Cahn–Hilliard–Cook equation (1.1).

The analysis of numerical methods for SPDEs without a global monotonicity assumption is even less explored. Pathwise convergence of a spectral Galerkin method for the stochastic Burgers equation is studied in [2, 3], while for the same equation convergence in probability is established in [33] for the Backward Euler method. The stochastic Navier–Stokes equation is considered in [5, 7], in particular, in [7] the authors obtain a result similar to our Theorem 5.4 (stated in a slightly different form). For the Cahn–Hilliard–Cook equation, convergence in probability is established for a finite difference scheme in [6]. In [21] strong convergence with rates is established for the spatial spectral Galerkin approximation for (1.1) and the stochastic Burgers equation driven by trace class noise in spatial dimension $d = 1$. Strong convergence of the finite element method without rate and without time discretization, is proved in [27, 28] under stronger assumptions both on the noise and the initial data than in the present paper. Therefore, this work can be viewed as the (non-trivial) extension of [27, 28] to a strongly convergent fully discrete scheme, still without rate but with improvements on the regularity requirement on the data. Finally, we mention the recent work [22], where strong convergence is proved, without rate, for a spectral nonlinearity-truncated accelerated exponential Euler-type approximation for the stochastic Kuramoto–Sivashinsky equation driven by space-time white noise in spatial dimension $d = 1$, an equation rather similar in structure to the Cahn–Hilliard–Cook equation.

The paper is organized as follows. In Section 2 we collect some background material from stochastic and functional analysis and introduce the finite element method in Subsection 2.2. In Section 3 some known results on the existence, uniqueness and regularity on the solution of the (1.1) are recalled. Section 4 contains moment bounds for the numerical solution, in particular, it contains the main technical result of the paper, Theorem 4.3. In Section 5 we prove a new error estimate for the derivative of the error in the spatial semidiscretization of the linear deterministic Cahn–Hilliard equation, (5.5) in Lemma 5.2. Then we proceed to prove a pathwise error bound in Theorem 5.4 and the main strong convergence result Theorem 5.5.

2. PRELIMINARIES

2.1. Norms and operators. Throughout the paper we will use various norms for linear operators on a Hilbert space H . We denote by $\mathcal{L}(H)$, the space of bounded linear operators on H with the usual operator norm denoted by $\|\cdot\|$. If for a positive semidefinite operator $T: H \rightarrow H$, the sum

$$\mathrm{Tr} T := \sum_{k=1}^{\infty} \langle T e_k, e_k \rangle < \infty$$

for an orthonormal basis (ONB) $\{e_k\}_{k \in \mathbb{N}}$ of H , then we say that T is trace class. In this case $\mathrm{Tr} T$, the trace of T , is independent of the choice of the ONB. If for an operator $T: H \rightarrow H$, the sum

$$\|T\|_{\mathrm{HS}}^2 := \sum_{k=1}^{\infty} \|T e_k\|^2 < \infty$$

for an ONB $\{e_k\}_{k \in \mathbb{N}}$ of H , then we say that T is Hilbert–Schmidt and call $\|T\|_{\mathrm{HS}}$ the Hilbert–Schmidt norm of T . The Hilbert–Schmidt norm of T is independent of the choice of the ONB. We have the following well-known properties of the trace and Hilbert–Schmidt norms, see, for example, [11, Appendix C],

$$(2.1) \quad \|T\| \leq \|T\|_{\mathrm{HS}}, \quad \|TS\|_{\mathrm{HS}} \leq \|T\|_{\mathrm{HS}} \|S\|, \quad \|ST\|_{\mathrm{HS}} \leq \|S\| \|T\|_{\mathrm{HS}},$$

$$(2.2) \quad \mathrm{Tr} Q = \|Q^{\frac{1}{2}}\|_{\mathrm{HS}}^2 = \|T\|_{\mathrm{HS}}^2 = \|T^*\|_{\mathrm{HS}}^2, \quad \text{if } Q = TT^*.$$

Next we introduce spaces and norms associated with A . Let $\mathcal{D} \subset \mathbf{R}^d$, $d = 1, 2, 3$, be a bounded convex domain with polygonal boundary $\partial\mathcal{D}$. We denote by $\|\cdot\|_{L_p}$ the standard norm in $L_p(\mathcal{D})$. In particular, we define $H = L_2(\mathcal{D})$ with standard inner product $\langle \cdot, \cdot \rangle$ and norm $\|\cdot\| = \|\cdot\|_{L_2}$, and

$$\dot{H} = \left\{ v \in H : \int_{\mathcal{D}} v \, dx = 0 \right\}.$$

Let $P: H \rightarrow \dot{H}$ define the orthogonal projector. Then $(I - P)v = |\mathcal{D}|^{-1} \int_{\mathcal{D}} v \, dx$ is the average of v . We also denote by $H^k(\mathcal{D})$ the standard Sobolev space. We define $A = -\Delta$ with domain of definition

$$D(A) = \left\{ v \in H^2(\mathcal{D}) : \frac{\partial v}{\partial n} = 0 \text{ on } \partial\mathcal{D} \right\}.$$

Then A is a positive definite, self-adjoint, unbounded, linear operator on \dot{H} with compact inverse. When extended to H as $Av = APv$ it has an orthonormal eigenbasis $\{\varphi_j\}_{j=0}^{\infty}$ with corresponding eigenvalues $\{\lambda_j\}_{j=0}^{\infty}$ such that

$$0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_j \leq \cdots, \quad \lambda_j \rightarrow \infty.$$

The first eigenfunction is constant, $\varphi_0 = |\mathcal{D}|^{-\frac{1}{2}}$.

We define

$$(2.3) \quad |v|_{\alpha} = \left(\sum_{j=1}^{\infty} \lambda_j^{\alpha} |\langle v, \varphi_j \rangle|^2 \right)^{\frac{1}{2}}, \quad \langle v, w \rangle_{\alpha} = \sum_{j=1}^{\infty} \lambda_j^{\alpha} \langle v, \varphi_j \rangle \langle w, \varphi_j \rangle, \quad \alpha \in \mathbf{R},$$

$$(2.4) \quad \|v\|_{\alpha} = (|v|_{\alpha}^2 + |\langle v, \varphi_0 \rangle|^2)^{\frac{1}{2}}, \quad \alpha \geq 0,$$

and corresponding spaces, for $\alpha \geq 0$,

$$\dot{H}^{\alpha} = D(A^{\frac{\alpha}{2}}) = \left\{ v \in \dot{H} : |v|_{\alpha} < \infty \right\}, \quad H^{\alpha} = \left\{ v \in H : \|v\|_{\alpha} < \infty \right\}.$$

For negative order $-\alpha < 0$ we define $\dot{H}^{-\alpha}$ by taking the closure of \dot{H} with respect to $|\cdot|_{-\alpha}$. For integer order $\alpha = k = 1, 2$, H^k coincide with the standard Sobolev spaces $H^k(\mathcal{D})$ with $\|\cdot\|_k$ equivalent to the standard norm $\|\cdot\|_{H^k(\mathcal{D})}$. For example,

$$\begin{aligned}\|v\|_1^2 &= |v|_1^2 + |\langle v, \varphi_0 \rangle|^2 = \|\nabla v\|^2 + |\langle v, \varphi_0 \rangle|^2, \\ \|v\|_2^2 &= |v|_2^2 + |\langle v, \varphi_0 \rangle|^2 = \|\Delta v\|^2 + |\langle v, \varphi_0 \rangle|^2\end{aligned}$$

are equivalent to the standard norms $\|v\|_{H^k(\mathcal{D})}^2$, $k = 1, 2$, by the Poincaré inequality and the regularity estimate for the elliptic Neumann problem.

We recall the fact the operator $-A^2$ is the infinitesimal generator of an analytic semigroup $E(t) = e^{-tA^2}$ on H ,

$$\begin{aligned}(2.5) \quad E(t) &= e^{-tA^2} v = \sum_{j=0}^{\infty} e^{-t\lambda_j^2} \langle v, \varphi_j \rangle \varphi_j = \sum_{j=1}^{\infty} e^{-t\lambda_j^2} \langle v, \varphi_j \rangle \varphi_j + \langle v, \varphi_0 \rangle \varphi_0 \\ &= e^{-tA^2} P v + (I - P)v.\end{aligned}$$

It follows that

$$(2.6) \quad \|A^\alpha E(t)v\| \leq \|v\|, \quad \leq C t^{-\frac{\alpha}{2}} e^{-ct} \|v\|, \quad v \in H, \quad \alpha > 0.$$

and

$$(2.7) \quad \left(\int_0^t s^{2j} \|A^{2j+1} E(s)v\|^2 ds \right)^{1/2} \leq C \|v\|, \quad v \in H, \quad j = 0, 1, 2, \dots$$

2.2. The finite element method. Let $\{\mathcal{T}_h\}_{h>0}$ denote a family of regular triangulations of \mathcal{D} with maximal mesh size h . Let S_h be the space of continuous functions on \mathcal{D} , which are piecewise polynomials of degree ≤ 1 with respect to \mathcal{T}_h . Hence, $S_h \subset H^1$. We also define $\dot{S}_h = P S_h$; that is,

$$\dot{S}_h = \left\{ v_h \in S_h : \int_{\mathcal{D}} v_h dx = 0 \right\}.$$

The space \dot{S}_h is introduced only for the purpose of theory but not for computation. Now we define the "discrete Laplacian" $A_h : S_h \rightarrow S_h$ by

$$\langle A_h v_h, w_h \rangle = \langle \nabla v_h, \nabla w_h \rangle, \quad \forall v_h \in S_h, w_h \in S_h.$$

The operator A_h is self-adjoint, positive definite on \dot{S}_h , positive semidefinite on S_h , and A_h has an orthonormal eigenbasis $\{\varphi_{h,j}\}_{j=0}^{N_h}$ with corresponding eigenvalues $\{\lambda_{h,j}\}_{j=0}^{N_h}$. We have

$$0 = \lambda_{h,0} < \lambda_{h,1} \leq \dots \leq \lambda_{h,j} \leq \dots \leq \lambda_{h,N_h}$$

and $\varphi_{h,0} = \varphi_0 = |\mathcal{D}|^{-\frac{1}{2}}$. Moreover, we define $E_h(t) = e^{-tA_h^2} : S_h \rightarrow S_h$ by

$$E_h(t) = e^{-tA_h^2} v_h = \sum_{j=0}^{N_h} e^{-t\lambda_{h,j}^2} \langle v_h, \varphi_{h,j} \rangle \varphi_{h,j} = \sum_{j=1}^{N_h} e^{-t\lambda_{h,j}^2} \langle v_h, \varphi_{h,j} \rangle \varphi_{h,j} + (I - P)v_h$$

and the orthogonal projector $P_h : H \rightarrow S_h$ by

$$(2.8) \quad \langle P_h v, w_h \rangle = \langle v, w_h \rangle \quad \forall v \in H, w_h \in S_h.$$

Clearly, $P_h : \dot{H} \rightarrow \dot{S}_h$ and

$$E_h(t) P_h v = E_h(t) P_h P v + (I - P)v.$$

We have a discrete analog of (2.6),

$$(2.9) \quad \|A_h^\alpha E_h(t) P_h v\| \leq C t^{-\frac{\alpha}{2}} e^{-ct} \|v\|, \quad v \in H, \alpha > 0.$$

Finally, we define the Ritz projector $R_h: \dot{H}^1 \rightarrow \dot{S}_h$ by

$$\langle \nabla R_h v, \nabla w_h \rangle = \langle \nabla v, \nabla w_h \rangle, \quad \forall v \in \dot{H}^1, w_h \in \dot{S}_h.$$

We extend it as $R_h: H^1 \rightarrow S_h$ by

$$(2.10) \quad R_h v = R_h P v + (I - P)v, \quad v \in H^1.$$

We then have the following bound for $R_h v - v = (R_h - I)Pv$ (cf. [34, Chapt. 1])

$$(2.11) \quad \|(R_h - I)v\| \leq C h^2 \|Av\|, \quad v \in \dot{H}^2.$$

Finally, we define discrete versions of the norms $|\cdot|_\alpha$ on \dot{H}^α :

$$(2.12) \quad |v_h|_{\alpha,h} = \|A_h^{\alpha/2} v_h\| = \left(\sum_{j=1}^{N_h} \lambda_{j,h}^\alpha |\langle v, \varphi_{j,h} \rangle|^2 \right)^{\frac{1}{2}}, \quad v_h \in \dot{S}_h, \alpha \in \mathbf{R}.$$

These are norms on \dot{S}_h and the corresponding scalar products are denoted $\langle \cdot, \cdot \rangle_{\alpha,h}$. We note that

$$(2.13) \quad |v_h|_1 = \|A_h^{\frac{1}{2}} v_h\| = \|\nabla v_h\| = \|A_h^{\frac{1}{2}} v_h\| = |v_h|_{1,h}, \quad v_h \in \dot{S}_h.$$

We assume that P_h is bounded with respect to the \dot{H}^1 norm

$$(2.14) \quad |P_h v|_1 \leq C |v|_1, \quad v \in \dot{H}^1.$$

This holds, for example, if the mesh family $\{\mathcal{T}_h\}_{h>0}$ is quasi-uniform. By combining this with (2.13), we obtain

$$(2.15) \quad \|A_h^{1/2} P_h v\| = |P_h v|_1 \leq C |v|_1 = C \|A^{1/2} v\|.$$

2.3. Various useful inequalities. We will also use the Burkholder–Davies–Gundy inequality for Itô-integrals of the form $\int_0^t \langle \eta(s), d\tilde{W}(s) \rangle$, where \tilde{W} is a \tilde{Q} -Wiener process. For this kind of integral, the Burkholder–Davies–Gundy inequality, [11, Lemma 7.2], takes the form

$$(2.16) \quad \mathbf{E} \sup_{t \in [0, T]} \left| \int_0^t \langle \eta(s), d\tilde{W}(s) \rangle \right|^p \leq C_p \mathbf{E} \left(\int_0^T \|\tilde{Q}^{\frac{1}{2}} \eta(s)\|^2 ds \right)^{\frac{p}{2}}, \quad p \geq 2.$$

Also, if Y is an H -valued Gaussian random variable with covariance operator \tilde{Q} , then, by [11, Corollary 2.17], we can bound its p -th moments via its covariance operator as

$$(2.17) \quad \mathbf{E} \|Y\|^{2p} \leq C_p (\mathbf{E} \|Y\|^2)^p = C_p (\text{Tr } \tilde{Q})^p = \|\tilde{Q}^{\frac{1}{2}}\|_{\text{HS}}^{2p}, \quad p \geq 1.$$

We will repeatedly apply this to the Itô integral $\int_s^t R dW(r)$, where R is a constant, possibly unbounded, operator on H and W is a \tilde{Q} -Wiener process. Then

$$\int_s^t R dW(r) = R(W(t) - W(s))$$

and (2.17) reads

$$(2.18) \quad \mathbf{E} \left\| \int_s^t R dW(r) \right\|^{2p} \leq C (t-s)^p \|R \tilde{Q}^{1/2}\|_{\text{HS}}^{2p}.$$

If $p = 1$, the inequality in (2.18) becomes an equality with $C = 1$. The inequality

$$(2.19) \quad \left| \sum_{j=K}^M a_j \right|^p \leq |M - K + 1|^{p-1} \sum_{j=K}^M |a_j|^p, \quad p \geq 1,$$

will be frequently utilised; it is a direct consequence of Hölder's inequality.

3. EXISTENCE, UNIQUENESS AND REGULARITY

Existence, uniqueness, and regularity of solutions to (1.1) has been studied in [10] with some minor improvements in [27]. Note that here we assume that X_0 is deterministic and that $X_0 \in \dot{H}$, so that $X(t) \in \dot{H}$. We summarize the results:

Theorem 3.1. *If $\|A^{\frac{1}{2}}Q^{\frac{1}{2}}\|_{\text{HS}} < \infty$ and $|X_0|_1 < \infty$, then there is a unique weak solution X of (1.1). Furthermore, there is $C_T > 0$ such that*

$$(3.1) \quad \mathbb{E} \sup_{t \in [0, T]} |X(t)|_1^2 + \mathbb{E} \sup_{t \in [0, T]} \|X(t)\|_{L_4}^4 \leq C_T.$$

In addition, X is also a mild solution, that is, it solves

$$(3.2) \quad X(t) = E(t)X_0 - \int_0^t AE(t-s)f(X(s)) \, ds + \int_0^t E(t-s) \, dW(s),$$

almost surely.

We also have pathwise Hölder regularity in time:

Proposition 3.2. *Let $\|A^{\frac{1}{2}}Q^{\frac{1}{2}}\|_{\text{HS}} < \infty$ and $|X_0|_1 < \infty$. Then, for all $\gamma \in [0, \frac{1}{2})$, there is a bounded nonnegative random variable K such that, almost surely,*

$$\sup_{t \neq s \in [0, T]} \frac{\|X(t) - X(s)\|}{|t - s|^\gamma} \leq K.$$

We omit the proof as it is analogous to the proof of [25, Proposition 3.2].

4. MOMENT BOUNDS

We start by proving a preliminary moment bound which will be used later on in a bootstrapping argument. We recall our assumption that $X_0 \in \dot{H}$, so that $X_h^0 = P_h X_0 \in \dot{S}_h$ and hence $X_h^j \in \dot{S}_h$ for $0 \leq j \leq N$.

Lemma 4.1. *Let $p \geq 1$. If $\|A_h^{-1/2}P_h Q^{1/2}\|_{\text{HS}} \leq K$ and $|X_h^0|_{-1,h} \leq L$, then there exist $C = C(p, T, K, L)$ and $k_0 > 0$ such that, for $0 < k < k_0$,*

$$(4.1) \quad \mathbb{E} \left(\sup_{1 \leq j \leq N} |X_h^j|_{-1,h}^{2p} \right) \leq C,$$

$$(4.2) \quad \mathbb{E} \left(\sum_{j=1}^N (|X_h^j - X_h^{j-1}|_{-1,h}^2 + k|X_h^j|_1^2) \right)^p \leq C.$$

Proof. Since $X_h^j \in \dot{S}_h$, we may multiply by $A_h^{-1}X_h^j$ in (1.6) to get

$$\begin{aligned} & \frac{1}{2} \left(|X_h^j|_{-1,h}^2 - |X_h^{j-1}|_{-1,h}^2 + |X_h^j - X_h^{j-1}|_{-1,h}^2 \right) + k|X_h^j|_1^2 + k \langle f(X_h^j), X_h^j \rangle \\ & = \langle P_h \Delta W^j, A_h^{-1}X_h^j \rangle, \end{aligned}$$

where we have used the self-adjointness of A_h , (2.13), and the identity

$$(4.3) \quad \langle X - Y, X \rangle_{-1,h} = \frac{1}{2}(|X|_{-1,h}^2 - |Y|_{-1,h}^2 + |X - Y|_{-1,h}^2).$$

Also, $\langle f(X_h^j), X_h^j \rangle \geq -C_0 - C_1 \|X_h^j\|^2$ by (1.3) and $\|X_h^j\|^2 \leq C_\epsilon |X_h^j|_{-1,h}^2 + \epsilon |X_h^j|_1^2$ for $\epsilon > 0$. Therefore,

$$\begin{aligned} & \frac{1}{2} \left(|X_h^j|_{-1,h}^2 - |X_h^{j-1}|_{-1,h}^2 + |X_h^j - X_h^{j-1}|_{-1,h}^2 \right) + ck |X_h^j|_1^2 \\ & \leq Ck + Ck |X_h^j|_{-1,h}^2 + \langle P_h \Delta W^j, X_h^j - X_h^{j-1} \rangle_{-1,h} + \langle P_h \Delta W^j, X_h^{j-1} \rangle_{-1,h}. \end{aligned}$$

Furthermore, $\langle P_h \Delta W^j, X_h^j - X_h^{j-1} \rangle_{-1,h} \leq C_\epsilon |P_h \Delta W^j|_{-1,h}^2 + \epsilon |X_h^j - X_h^{j-1}|_{-1,h}^2$. Hence,

$$(4.4) \quad \begin{aligned} & \frac{1}{2} \left(|X_h^j|_{-1,h}^2 - |X_h^{j-1}|_{-1,h}^2 \right) + c |X_h^j - X_h^{j-1}|_{-1,h}^2 + ck |X_h^j|_1^2 \\ & \leq C \left(k + k |X_h^j|_{-1,h}^2 + |P_h \Delta W^j|_{-1,h}^2 + \langle P_h \Delta W^j, X_h^{j-1} \rangle_{-1,h} \right). \end{aligned}$$

Summing with respect to j in (4.4), thus yields

$$(4.5) \quad \begin{aligned} & |X_h^n|_{-1,h}^2 + \sum_{j=1}^n \left(|X_h^j - X_h^{j-1}|_{-1,h}^2 + k |X_h^j|_1^2 \right) \leq C \left(T + |X_h^0|_{-1,h}^2 \right. \\ & \quad \left. + \sum_{j=1}^n k |X_h^j|_{-1,h}^2 + \sum_{j=1}^n |P_h \Delta W^j|_{-1,h}^2 + \sum_{j=1}^n \langle P_h \Delta W^j, X_h^{j-1} \rangle_{-1,h} \right). \end{aligned}$$

We drop the sum on the left, take the p 'th power and the supremum with respect to n , and then the expectation, using also (2.19) repeatedly, to get

$$(4.6) \quad \begin{aligned} & \mathbf{E} \sup_{1 \leq n \leq N} |X_h^n|_{-1,h}^{2p} \leq \mathbf{E} \sup_{1 \leq n \leq N} C \left\{ T + |X_h^0|_{-1,h}^2 + \sum_{j=1}^n k |X_h^j|_{-1,h}^2 \right. \\ & \quad \left. + \sum_{j=1}^n |P_h \Delta W^j|_{-1,h}^2 + \sum_{j=1}^n \langle P_h \Delta W^j, X_h^{j-1} \rangle_{-1,h} \right\}^p \\ & \leq C \mathbf{E} \left\{ T^p + |X_h^0|_{-1,h}^{2p} + \left(\sum_{j=1}^N k |X_h^j|_{-1,h}^2 \right)^p \right. \\ & \quad \left. + \left(\sum_{j=1}^N |P_h \Delta W^j|_{-1,h}^2 \right)^p + \sup_{1 \leq n \leq N} \left(\sum_{j=1}^n \langle P_h \Delta W^j, X_h^{j-1} \rangle_{-1,h} \right)^p \right\} \\ & \leq C \left\{ T^p + |X_h^0|_{-1,h}^{2p} + (Nk)^{p-1} \sum_{j=1}^N k \mathbf{E} |X_h^j|_{-1,h}^{2p} \right. \\ & \quad \left. + N^{p-1} \sum_{j=1}^N \mathbf{E} |P_h \Delta W^j|_{-1,h}^{2p} + \mathbf{E} \sup_{1 \leq n \leq N} \left(\sum_{j=1}^n \langle P_h \Delta W^j, X_h^{j-1} \rangle_{-1,h} \right)^p \right\}. \end{aligned}$$

By (2.17), we have

$$(4.7) \quad \mathbf{E} |P_h \Delta W^j|_{-1,h}^{2p} \leq C_p (\mathbf{E} |P_h \Delta W^j|_{-1,h}^2)^p \leq C_p k^p \|A_h^{-1/2} P_h Q^{1/2}\|_{\text{HS}}^{2p}$$

and, by Cauchy's inequality and (2.16),

$$\begin{aligned}
& \mathbf{E} \sup_{1 \leq n \leq N} \left(\sum_{j=1}^n \langle P_h \Delta W^j, X_h^{j-1} \rangle_{-1,h} \right)^p = \mathbf{E} \sup_{1 \leq n \leq N} \left(\sum_{j=1}^n \langle \Delta W^j, A_h^{-1} X_h^{j-1} \rangle \right)^p \\
& \leq C \mathbf{E} \left(\sum_{j=1}^N k \|Q^{1/2} A_h^{-1} X_h^{j-1}\|^2 \right)^{p/2} \\
& \leq C \left\{ 1 + \mathbf{E} \left(\sum_{j=1}^N k \|Q^{1/2} A_h^{-1} X_h^{j-1}\|^2 \right)^p \right\} \\
(4.8) \quad & \leq C \left\{ 1 + (Nk)^{p-1} \sum_{j=1}^N k \mathbf{E} \|Q^{1/2} A_h^{-1} P_h X_h^{j-1}\|^{2p} \right\} \\
& \leq C \left\{ 1 + (Nk)^{p-1} \sum_{j=1}^N k \|Q^{1/2} A_h^{-1/2} P_h\|^{2p} \mathbf{E} |X_h^{j-1}|_{-1,h}^{2p} \right\} \\
& = C \left\{ 1 + (Nk)^{p-1} \sum_{j=1}^N k \|A_h^{-1/2} P_h Q^{1/2}\|^{2p} \mathbf{E} |X_h^{j-1}|_{-1,h}^{2p} \right\}.
\end{aligned}$$

As $Nk = T$, $\|A_h^{-1/2} P_h Q^{1/2}\| \leq \|A_h^{-1/2} P_h Q^{1/2}\|_{\text{HS}} \leq K$, and $|X_h^0|_{-1,h} \leq L$, by inserting (4.7) and (4.8) into (4.6), we see that

$$\begin{aligned}
(4.9) \quad & \mathbf{E} \sup_{1 \leq n \leq N} |X_h^n|_{-1,h}^{2p} \leq C(p, T, K, L) \left(1 + \sum_{j=1}^{N-1} k \mathbf{E} |X_h^j|_{-1,h}^{2p} \right) \\
& \leq C(p, T, K, L) \left(1 + \sum_{n=1}^{N-1} k \mathbf{E} \sup_{1 \leq j \leq n} \|X_h^j\|_{-1,h}^{2p} \right).
\end{aligned}$$

The first part of the statement thus follows from Gronwall's lemma for k small enough. Having this result at hand one may return to (4.6) to prove the rest of the statement by a similar procedure. \square

Lemma 4.2. *Suppose that $\|A^{1/2} Q^{1/2}\|_{\text{HS}} \leq K$ and $|X_h^0|_{-1,h} \leq L$. Then, for every $\epsilon, \delta > 0$ and $p \geq 1$, there are $C_1 = C_1(T, \epsilon, \delta, p, K, L)$, $C_2 = C_2(T, p)$, and $k_0 = k_0(\epsilon, p)$ such that for $0 < k < k_0$,*

$$\begin{aligned}
(4.10) \quad & \mathbf{E} \left(\sup_{1 \leq j \leq N} \|X_h^j\|^{2p} \right) + \mathbf{E} \left(\sum_{j=1}^N \|X_h^j - X_h^{j-1}\|^2 \right)^p \\
& \leq C_1 + C_2 \delta \mathbf{E} \left(\sum_{j=1}^N k |A_h X_h^j + P_h f(X_h^j)|_1^2 \right)^{\frac{1+\epsilon}{2} p}.
\end{aligned}$$

Proof. By taking inner products with $X_h^j \in \dot{S}_h$ in (1.6) we get

$$\begin{aligned}
& \frac{1}{2} \left(\|X_h^j\|^2 - \|X_h^{j-1}\|^2 + \|X_h^j - X_h^{j-1}\|^2 \right) \\
& \quad + k \langle A_h X_h^j + P_h f(X_h^j), X_h^j \rangle_1 = \langle \Delta W^j, X_h^j \rangle,
\end{aligned}$$

Summing with respect to j , using analogous arguments as in the previous proof, thus yields

$$\begin{aligned} \|X_h^n\|^2 + \sum_{j=1}^n \|X_h^j - X_h^{j-1}\|^2 &\leq C \left(\|X_h^0\|^2 + \sum_{j=1}^n k |\langle A_h X_h^j + P_h f(X_h^j), X_h^j \rangle_1| \right. \\ &\quad \left. + \sum_{j=1}^n \|\Delta W^j\|^2 + \sum_{j=1}^n \langle \Delta W^j, X_h^{j-1} \rangle \right). \end{aligned}$$

Therefore,

$$\begin{aligned} &\mathbf{E} \sup_{1 \leq j \leq N} \|X_h^j\|^{2p} + \mathbf{E} \left(\sum_{j=1}^N \|X_h^j - X_h^{j-1}\|^2 \right)^p \\ &\leq C \left\{ \|X_h^0\|^{2p} + \mathbf{E} \left(\sum_{j=1}^N k |\langle A_h X_h^j + P_h f(X_h^j), X_h^j \rangle_1| \right)^p \right. \\ &\quad \left. + \mathbf{E} \left(\sum_{j=1}^N \|\Delta W^j\|^2 \right)^p + \mathbf{E} \sup_{1 \leq n \leq N} \left(\sum_{j=1}^n \langle A_h^{1/2} P_h \Delta W^j, A_h^{-1/2} X_h^{j-1} \rangle \right)^p \right\}. \end{aligned}$$

The last term can be bounded, similarly to the previous proof, by $C_p T^p \|Q^{1/2}\|_{\text{HS}}^{2p}$. Using the bound (4.1) and a calculation similar to (4.8) we obtain that

$$\begin{aligned} \mathbf{E} \sup_{1 \leq n \leq N} \left(\sum_{j=1}^n \langle A_h^{1/2} P_h \Delta W^j, A_h^{-1/2} X_h^{j-1} \rangle \right)^p &\leq C \left(1 + T^p \|A_h^{1/2} P_h Q^{1/2}\|^{2p} \right) \\ &\leq C \left(1 + T^p \|A^{1/2} Q^{1/2}\|^{2p} \right), \end{aligned}$$

where we also used (2.15). Finally,

$$\begin{aligned} &\mathbf{E} \left(\sum_{j=1}^N k |\langle A_h X_h^j + P_h f(X_h^j), X_h^j \rangle_1| \right)^p \leq \mathbf{E} \left(\sum_{j=1}^N k |A_h X_h^j + P_h f(X_h^j)|_1 |X_h^j|_1 \right)^p \\ &\leq \mathbf{E} \left\{ \left(\sum_{j=1}^N k |A_h X_h^j + P_h f(X_h^j)|_1^2 \right)^{p/2} \left(\sum_{j=1}^N k |X_h^j|_1^2 \right)^{p/2} \right\} \\ &\leq \mathbf{E} \left\{ \delta \left(\sum_{j=1}^N k |A_h X_h^j + P_h f(X_h^j)|_1^2 \right)^{\frac{1+\epsilon}{2}p} + C_{\epsilon, \delta} \left(\sum_{j=1}^N k |X_h^j|_1^2 \right)^{\frac{1+\epsilon}{2\epsilon}p} \right\}, \end{aligned}$$

and the proof is complete in view of (4.2). \square

We next prove the main stability result of the paper. In the proof we denote by $P_\alpha(x)$, $x = (x_1, \dots, x_N)$, $x_i \geq 0$, a nonnegative expression such that $P_\alpha(x) \leq C(1 + \sum_{i=1}^N x_i^\alpha)$.

Theorem 4.3. *Let $p \geq 1$. If $\|A^{1/2} Q^{1/2}\|_{\text{HS}} \leq K$ and*

$$(4.11) \quad |X_h^0|_1 + \|X^0\| + |X_h^0|_{-1, h} + \mathcal{F}(X_h^0) + |A_h X_h^0 + P_h f(X_h^0)|_1 \leq L,$$

then there exists $C, k_0 > 0$, depending on p, K, L , and T , such that for $0 < k < k_0$,

$$(4.12) \quad \mathbf{E} \sup_{1 \leq j \leq N} |X_h^j|_1^{2p} + \mathbf{E} \sup_{1 \leq n \leq N} \mathcal{F}(X_h^n)^p + \mathbf{E} \left(\sum_{j=1}^N k |A_h X_h^j + P_h f(X_h^j)|_1^2 \right)^p \leq C.$$

Proof. Note first that from (1.4) it follows for $X, Y \in \dot{H}^0$ that

$$\mathcal{F}(X) - \mathcal{F}(Y) \leq \langle f(X), X - Y \rangle + \frac{1}{2}\beta^2 \|X - Y\|^2.$$

Using this and the analog of (4.3) with the seminorm on \dot{H}^1 , taking $X = X_h^j$ and $Y = X_h^{j-1}$, we obtain

$$\begin{aligned} (4.13) \quad & \frac{1}{2}|X_h^j|_1^2 + \mathcal{F}(X_h^j) - \frac{1}{2}|X_h^{j-1}|_1^2 - \mathcal{F}(X_h^{j-1}) + \frac{1}{2}|X_h^j - X_h^{j-1}|_1^2 \\ & \leq \langle X_h^j, X_h^j - X_h^{j-1} \rangle_1 + \langle f(X_h^j), X_h^j - X_h^{j-1} \rangle + C\|X_h^j - X_h^{j-1}\|^2 \\ & = \langle A_h X_h^j + P_h f(X_h^j), X_h^j - X_h^{j-1} \rangle + C\|X_h^j - X_h^{j-1}\|^2. \end{aligned}$$

Invoking (1.6), we have here

$$\begin{aligned} & \langle A_h X_h^j + P_h f(X_h^j), X_h^j - X_h^{j-1} \rangle \\ & = \langle A_h X_h^j + P_h f(X_h^j), -k(A_h^2 X_h^j + A_h P_h f(X_h^j)) + P_h \Delta W^j \rangle \\ & = -k|A_h X_h^j + P_h f(X_h^j)|_1^2 + \langle A_h X_h^j + P_h f(X_h^j), P_h \Delta W^j \rangle \\ & = -k|A_h X_h^j + P_h f(X_h^j)|_1^2 + \langle A_h(X_h^j - X_h^{j-1}), P_h \Delta W^j \rangle \\ & \quad + \langle f(X_h^j) - f(X_h^{j-1}), P_h \Delta W^j \rangle + \langle A_h X_h^{j-1} + P_h f(X_h^{j-1}), P_h \Delta W^j \rangle. \end{aligned}$$

Now, by (2.15),

$$(4.14) \quad \langle A_h(X_h^j - X_h^{j-1}), P_h \Delta W^j \rangle \leq \epsilon |X_h^j - X_h^{j-1}|_1^2 + C_\epsilon |\Delta W^j|_1^2$$

and, as $P_h \Delta W^j \in \dot{S}_h \subset \dot{H}$,

$$\begin{aligned} (4.15) \quad & |\langle f(X_h^j) - f(X_h^{j-1}), P_h \Delta W^j \rangle| = |\langle P(f(X_h^j) - f(X_h^{j-1})), P_h \Delta W^j \rangle| \\ & = |\langle A_h^{-1/2} P_h P(f(X_h^j) - f(X_h^{j-1})), A_h^{1/2} P_h \Delta W^j \rangle| \\ & \leq \|A_h^{-1/2} P_h P(f(X_h^j) - f(X_h^{j-1}))\| |P_h \Delta W^j|_1 \\ & \leq C \|A_h^{-1/2} P_h P(f(X_h^j) - f(X_h^{j-1}))\| |\Delta W^j|_1. \end{aligned}$$

Furthermore, (1.5) implies, using also (2.18) from Lemma 2.5 in [27] in the first inequality below, that

$$\begin{aligned} (4.16) \quad & \|A_h^{-1/2} P_h P(f(X_h^j) - f(X_h^{j-1}))\| \leq C \|f(X_h^j) - f(X_h^{j-1})\|_{L_{6/5}(\mathcal{D})} \\ & \leq C \left(\int_{\mathcal{D}} |X_h^j - X_h^{j-1}|^{6/5} (1 + (X_h^j)^2 + (X_h^{j-1})^2)^{6/5} d\xi \right)^{5/6} \\ & \leq C \left(\int_{\mathcal{D}} |X_h^j - X_h^{j-1}|^6 d\xi \right)^{1/6} \left(\int_{\mathcal{D}} (1 + (X_h^j)^2 + (X_h^{j-1})^2)^{3/2} d\xi \right)^{2/3} \\ & \leq C \|X_h^j - X_h^{j-1}\|_{L_6(\mathcal{D})} \left(1 + \|X_h^j\|_{L_3(\mathcal{D})}^2 + \|X_h^{j-1}\|_{L_3(\mathcal{D})}^2 \right). \end{aligned}$$

Further, with $0 < p < q < r$ we have that, for $\lambda = \frac{p}{q} \frac{r-q}{r-p}$,

$$(4.17) \quad \|X\|_{L_q(\mathcal{D})} \leq \|X\|_{L_p(\mathcal{D})}^\lambda \|X\|_{L_r(\mathcal{D})}^{1-\lambda},$$

if $X \in L_p(\mathcal{D}) \cap L_r(\mathcal{D})$ (see [13, Proposition 6.10]). We may take $p = 2$, $q = 3$, $r = 4$, and hence $\lambda = \frac{1}{3}$, to conclude that

$$(4.18) \quad \|X_h^j\|_{L_3(\mathcal{D})}^2 \leq \|X_h^j\|_{L_2(\mathcal{D})}^{2/3} \|X_h^j\|_{L_4(\mathcal{D})}^{4/3}.$$

Thus, from (4.15), (4.16) and (4.18), as also, by Sobolev's inequality and the fact that $X_h^j \in \dot{S}_h \subset \dot{H}^1$, we have $\|X_h^j - X_h^{j-1}\|_{L_6(\mathcal{D})} \leq C|X_h^j - X_h^{j-1}|_1$, it follows that

$$\begin{aligned}
 (4.19) \quad & |\langle f(X_h^j) - f(X_h^{j-1}), P_h \Delta W^j \rangle| \leq C|\Delta W^j|_1 |X_h^j - X_h^{j-1}|_1 \\
 & \times P_{2/3}(\|X_h^j\|_{L_2(\mathcal{D})}, \|X_h^{j-1}\|_{L_2(\mathcal{D})}) P_{4/3}(\|X_h^j\|_{L_4(\mathcal{D})}, \|X_h^{j-1}\|_{L_4(\mathcal{D})}) \\
 & \leq \epsilon |X_h^j - X_h^{j-1}|_1^2 + C|\Delta W^j|_1^2 P_{4/3}(\|X_h^j\|_{L_2(\mathcal{D})}, \|X_h^{j-1}\|_{L_2(\mathcal{D})}) \\
 & \times P_{8/3}(\|X_h^j\|_{L_4(\mathcal{D})}, \|X_h^{j-1}\|_{L_4(\mathcal{D})}).
 \end{aligned}$$

Thus with $0 < \epsilon < \frac{1}{4}$ we get after inserting (4.14) and (4.19) into (4.13) and rearranging that

$$\begin{aligned}
 (4.20) \quad & \frac{1}{2}|X_h^j|_1^2 + \mathcal{F}(X_h^j) - \frac{1}{2}|X_h^{j-1}|_1^2 - \mathcal{F}(X_h^{j-1}) + c|X_h^j - X_h^{j-1}|_1^2 + k|A_h X_h^j + P_h f(X_h^j)|_1^2 \\
 & \leq C_\epsilon \left(|\Delta W^j|_1^2 P_{4/3}(\|X_h^j\|_{L_2(\mathcal{D})}, \|X_h^{j-1}\|_{L_2(\mathcal{D})}) P_{8/3}(\|X_h^j\|_{L_4(\mathcal{D})}, \|X_h^{j-1}\|_{L_4(\mathcal{D})}) \right. \\
 & \quad \left. + |\Delta W^j|_1^2 \right) + \langle A_h X_h^{j-1} + P_h f(X_h^{j-1}), \Delta W^j \rangle + C\|X_h^j - X_h^{j-1}\|^2.
 \end{aligned}$$

Summing with respect to j then yields

$$\begin{aligned}
 (4.21) \quad & \frac{1}{2}|X_h^n|_1^2 + \mathcal{F}(X_h^n) + \sum_{j=1}^n \left(|X_h^j - X_h^{j-1}|_1^2 + k|A_h X_h^j + P_h f(X_h^j)|_1^2 \right) \\
 & \leq C \left(|X_h^0|_1^2 + \mathcal{F}(X_h^0) + \sum_{j=1}^n \|X_h^j - X_h^{j-1}\|^2 + \sum_{j=1}^n |\Delta W^j|_1^2 \right. \\
 & \quad \left. + \sum_{j=1}^n |\Delta W^j|_1^2 P_{4/3}(\|X_h^j\|_{L_2(\mathcal{D})}, \|X_h^{j-1}\|_{L_2(\mathcal{D})}) P_{8/3}(\|X_h^j\|_{L_4(\mathcal{D})}, \|X_h^{j-1}\|_{L_4(\mathcal{D})}) \right) \\
 & \quad + \sum_{j=1}^n \langle A_h X_h^{j-1} + P_h f(X_h^{j-1}), \Delta W^j \rangle.
 \end{aligned}$$

It follows in a similar way as in the proof of Lemma 4.1, using also (4.10) with $\epsilon = 1$ and $\delta > 0$ small enough for the third term on the right hand side above so that it can be absorbed into the third term in the left hand side below, that

$$\begin{aligned}
 (4.22) \quad & \mathbf{E} \sup_{1 \leq n \leq N} |X_h^n|_1^{2p} + \mathbf{E} \sup_{1 \leq n \leq N} \mathcal{F}(X_h^n)^p + \mathbf{E} \left(\sum_{j=1}^N k|A_h X_h^j + P_h f(X_h^j)|_1^2 \right)^p \\
 & \leq C(p, T) \left(1 + |X_h^0|_1^{2p} + \mathcal{F}(X_h^0)^p + \mathbf{E} \left(\sum_{j=1}^N |\Delta W^j|_1^2 \right)^p \right. \\
 & \quad \left. + N^{p-1} \mathbf{E} \sum_{j=1}^N \left(|\Delta W^j|_1^{2p} P_{4p/3}(\|X_h^j\|_{L_2(\mathcal{D})}, \|X_h^{j-1}\|_{L_2(\mathcal{D})}) \right. \right. \\
 & \quad \left. \left. \times P_{8p/3}(\|X_h^j\|_{L_4(\mathcal{D})}, \|X_h^{j-1}\|_{L_4(\mathcal{D})}) \right) \right) \\
 & \quad + \mathbf{E} \sup_{1 \leq n \leq N} \left| \sum_{j=1}^n \langle A_h X_h^{j-1} + P_h f(X_h^{j-1}), \Delta W^j \rangle \right|^p.
 \end{aligned}$$

The first two terms to the right of the inequality are bounded by assumption. The third term may be bounded in a similar way as the corresponding term in Lemma 4.1 using now instead of (4.7), that $\mathbf{E}|\Delta W^j|_1^{2p} \leq k^p C \|A^{1/2} Q^{1/2}\|_{\text{HS}}^{2p}$. To bound the fourth term we invoke Hölder's inequality, first with conjugate exponents $q_1, q'_1 > 1$ and then with $q_2, q'_2 > 1$, to get

$$\begin{aligned}
& N^{p-1} \mathbf{E} \sum_{j=1}^N (|\Delta W^j|_1^{2p} P_{4p/3}(\|X_h^j\|_{L_2(\mathcal{D})}, \|X_h^{j-1}\|_{L_2(\mathcal{D})}) \\
& \quad \times P_{8p/3}(\|X_h^j\|_{L_4(\mathcal{D})}, \|X_h^{j-1}\|_{L_4(\mathcal{D})})) \\
& \leq N^{p-1} \sum_{j=1}^N (\mathbf{E}|\Delta W^j|_1^{2pq'_1})^{1/q'_1} \left[\mathbf{E}(P_{4pq_1/3}(\|X_h^j\|_{L_2(\mathcal{D})}, \|X_h^{j-1}\|_{L_2(\mathcal{D})}) \right. \\
& \quad \times P_{8pq_1/3}(\|X_h^j\|_{L_4(\mathcal{D})}, \|X_h^{j-1}\|_{L_4(\mathcal{D})})) \Big]^{1/q_1} \\
& = C \|A^{1/2} Q^{1/2}\|_{\text{HS}}^{2p} k^{p-1} N^{p-1} \sum_{j=1}^n k \left[\mathbf{E}(P_{4pq_1/3}(\|X_h^j\|_{L_2(\mathcal{D})}, \|X_h^{j-1}\|_{L_2(\mathcal{D})}) \right. \\
& \quad \times P_{8pq_1/3}(\|X_h^j\|_{L_4(\mathcal{D})}, \|X_h^{j-1}\|_{L_4(\mathcal{D})})) \Big]^{1/q_1} \\
(4.23) \quad & \leq CT^{p-1} \sum_{j=1}^N k \left[\left(\mathbf{E}(P_{4pq_1q'_2/3}(\|X_h^j\|_{L_2(\mathcal{D})}, \|X_h^{j-1}\|_{L_2(\mathcal{D})})) \right)^{1/q'_2} \right. \\
& \quad \times \left. \left(\mathbf{E}(P_{8pq_1q_2/3}(\|X_h^j\|_{L_4(\mathcal{D})}, \|X_h^{j-1}\|_{L_4(\mathcal{D})})) \right)^{1/q_2} \right]^{1/q_1} \\
& \leq C_T \left\{ 1 + \sum_{j=1}^N k \left(\mathbf{E}(P_{4pq_1q'_2/3}(\|X_h^j\|_{L_2(\mathcal{D})}, \|X_h^{j-1}\|_{L_2(\mathcal{D})})) \right)^{1/q'_2} \right. \\
& \quad \times \left. \left(\mathbf{E}(P_{8pq_1q_2/3}(\|X_h^j\|_{L_4(\mathcal{D})}, \|X_h^{j-1}\|_{L_4(\mathcal{D})})) \right)^{1/q_2} \right\} \\
& \leq C_T \left\{ \left(1 + \sum_{j=1}^N k \left[\mathbf{E}(P_{4pq_1q'_2/3}(\|X_h^j\|_{L_2(\mathcal{D})}, \|X_h^{j-1}\|_{L_2(\mathcal{D})})) \right. \right. \right. \\
& \quad \left. \left. \left. + \mathbf{E}(P_{8pq_1q_2/3}(\|X_h^j\|_{L_4(\mathcal{D})}, \|X_h^{j-1}\|_{L_4(\mathcal{D})})) \right] \right) \right\}.
\end{aligned}$$

If $q_1 q_2 = 3/2$ and $q_1 q'_2 (1 + \epsilon)/3 = 1$, then using (1.2), we get

$$\begin{aligned}
(4.24) \quad & P_{8pq_1q_2/3}(\|X_h^j\|_{L_4(\mathcal{D})}, \|X_h^{j-1}\|_{L_4(\mathcal{D})}) = C(\|X_h^j\|_{L_4(\mathcal{D})}^{4p} + \|X_h^{j-1}\|_{L_4(\mathcal{D})}^{4p} + 1) \\
& \leq C(\mathcal{F}(X_h^j)^p + \mathcal{F}(X_h^{j-1})^p + 1).
\end{aligned}$$

Furthermore, it follows from Lemma 4.2 that, with $C_2 = C_2(T, p) > 0$ independent of $\delta > 0$,

$$\begin{aligned}
& \mathbf{E}(P_{4pq_1q'_2/3}(\|X_h^j\|_{L_2(\mathcal{D})}, \|X_h^{j-1}\|_{L_2(\mathcal{D})})) \\
&= C\mathbf{E}(\|X_h^j\|_{L_2(\mathcal{D})}^{4pq_1q'_2/3} + \|X_h^{j-1}\|_{L_2(\mathcal{D})}^{4pq_1q'_2/3} + 1) \\
(4.25) \quad &\leq C(1 + \|X_h^0\|_{L_2(\mathcal{D})}^{4p}) + C_2\delta\mathbf{E}\left(\sum_{j=1}^N k|A_hX_h^j + P_hf(X_h^j)|_1^2\right)^{\frac{pq_1q'_2(1+\epsilon)}{3}} \\
&= C(1 + \|X_h^0\|_{L_2(\mathcal{D})}^{4p}) + C_2\delta\mathbf{E}\left(\sum_{j=1}^N k|A_hX_h^j + P_hf(X_h^j)|_1^2\right)^p.
\end{aligned}$$

We take (this is just one possible set of admissible values) $\epsilon = 1/4$, $q_1 = 18/17$, $q_2 = 17/12$ and $q'_2 = 17/5$. Thus inserting (4.24) and (4.25) into (4.23) we get

$$\begin{aligned}
& N^{p-1}\mathbf{E}\sum_{j=1}^N |\Delta W^j|_1^{2p} P_{4p/3}(\|X_h^j\|_{L_2(\mathcal{D})}) P_{4p/3}(\|X_h^j\|_{L_4(\mathcal{D})}) \\
&\leq C\left(1 + \|X_h^0\|_{L_2(\mathcal{D})}^{4p} + \sum_{j=0}^N k\mathbf{E}\mathcal{F}(X_h^j)^p\right) + C_2\delta\mathbf{E}\left(\sum_{j=1}^N k|A_hX_h^j + P_hf(X_h^j)|_1^2\right)^p.
\end{aligned}$$

It remains to treat the Itô integral in (4.22). For this we invoke the Cauchy inequality and Burkholder–Davies–Gundy inequality to conclude that

$$\begin{aligned}
(4.27) \quad & \mathbf{E}\sup_{1 \leq n \leq N} \left| \sum_{j=1}^n \langle A_hX_h^{j-1} + P_hf(X_h^{j-1}), \Delta W^j \rangle \right|^p \\
&\leq C(1 + \epsilon'\mathbf{E}\sup_{0 \leq n \leq N} \left| \sum_{j=1}^n \langle A_hX_h^{j-1} + P_hf(X_h^{j-1}), \Delta W^j \rangle \right|^{2p}) \\
&\leq C(1 + \epsilon'\mathbf{E}\left(\sum_{j=1}^N k\|Q^{1/2}(A_hX_h^{j-1} + P_hf(X_h^{j-1}))\|^2\right)^p) \\
&\leq C(1 + \epsilon'\mathbf{E}\left(\sum_{j=1}^N k\|Q^{1/2}A^{-1/2}P\| \|(A_hX_h^{j-1} + P_hf(X_h^{j-1}))\|_1^2\right)^p)
\end{aligned}$$

Thus, since $\|Q^{1/2}A^{-1/2}P\| < \infty$, if we take $\epsilon' > 0$ small enough in (4.27) and $\delta > 0$ small enough in (4.26), from (4.22) we may conclude that

$$\begin{aligned}
(4.28) \quad & \mathbf{E}\sup_{1 \leq n \leq N} |X_h^n|_1^{2p} + \mathbf{E}\sup_{1 \leq n \leq N} \mathcal{F}(X_h^n)^p + \mathbf{E}\left(\sum_{j=1}^N k|A_hX_h^j + P_hf(X_h^j)|_1^2\right)^p \\
&\leq C\left(1 + |X_h^0|_1^{2p} + \|X_h^0\|_{L_2(\mathcal{D})}^{4p} + (1+k)\mathcal{F}(X_h^0)^p + k|A_hX_h^0 + P_hf(X_h^0)|_1^{2p}\right. \\
&\quad \left. + \sum_{j=1}^N k\mathbf{E}\sup_{1 \leq n \leq j} \mathcal{F}(X_h^n)^p\right).
\end{aligned}$$

So, if $Ck < 1$ the claimed result follows from Gronwall's lemma. \square

5. CONVERGENCE

Recall from Theorem 3.1 that X satisfies the mild formulation representation

$$(5.1) \quad X(t) = E(t)X_0 - \int_0^t E(t-s)Af(X(s))ds + \int_0^t E(t-s)dW(s).$$

Similarly, equation (1.6) has the mild formulation

$$(5.2) \quad X_h^n = R_{k,h}^n P_h X_0 - \sum_{j=1}^n R_{k,h}^{n-j+1} A_h P_h f(X_h^j) + \sum_{j=1}^n R_{k,h}^{n-j+1} P_h \Delta W^j.$$

where $R_{k,h}^n = (I + \Delta t A_h^2)^{-n}$. We first need a maximal type error estimate for the stochastic convolution. Define the backward Euler approximation of the stochastic convolution $W_A(t) := \int_0^t E(t-s)dW(s)$ by

$$W_{A_h}^n := \sum_{k=1}^n R_{k,h}^{n-k+1} P_h \Delta W^k = \sum_{k=1}^n \int_{t_{k-1}}^{t_k} R_{k,h}^{n-k+1} P_h dW(s).$$

Lemma 5.1. *Let $\epsilon \in (0, \frac{1}{2}]$, $\beta \in [1, 2]$, and $p > \frac{2}{\epsilon}$. Then there is $C = C(p, \epsilon, T)$ such that*

$$\left(\mathbf{E} \left(\sup_{t_n \in [0, T]} \|W_A(t_n) - W_{A_h}^n\|^p \right) \right)^{1/p} \leq C(h^\beta + k^{\beta/4}) \|A^{(\beta-2)/2+\epsilon} P Q^{1/2}\|_{\text{HS}}.$$

Proof. The proof is completely analogous to the proof of [25, Proposition 5.1] using a discrete factorization method using the deterministic error estimates

$$(5.3) \quad \begin{aligned} \|(E(t_n) - R_{k,h}^n)P_h v\| &\leq C(h^\beta + k^{\beta/4})|v|_\beta, \quad t_n \geq 0, \\ \|(E(t_n) - R_{k,h}^n)P_h v\| &\leq C(h^\beta + k^{\beta/4})t_n^{-(\beta-\gamma)/4}|v|_\gamma, \quad t > 0, \gamma \in [-1, 1]. \end{aligned}$$

from [12, Lemma 5.5] and [32, Theorem 2.2] and the analyticity of the semigroup E . \square

Lemma 5.2. *The following deterministic error estimates hold for $v \in H$:*

$$(5.4) \quad \|A_h R_{k,h}^n P_h v - A_h E_h(t_n) P_h v\| \leq C k^{1/2} t^{-1} \|v\|,$$

$$(5.5) \quad \|A_h E_h(t) P_h v - A E(t) v\| \leq C h^2 t^{-1} \|v\|.$$

Proof. Note that it is enough to consider $v \in \dot{H}$ as for v constant, the above differences equal 0. Estimate (5.4) follows by a simple spectral argument. For estimate (5.5), we first write, for $v \in \dot{H}$,

$$\begin{aligned} \|A_h E_h(t) P_h v - A E(t) v\| &\leq \|(A_h^2 E_h(t) P_h - A^2 E(t)) A^{-1} v\| \\ &\quad + \|A_h^2 E_h(t) P_h (A_h^{-1} P_h - A^{-1}) v\| \\ &= \|D_t(E_h(t) P_h - E(t)) A^{-1} v\| + \|D_t E_h(t) P_h (A_h^{-1} P_h - A^{-1}) v\|. \end{aligned}$$

The desired bound for the last term follows immediately from (2.11) and (2.9). The first term is an error estimate for the time derivative of the solution of the linear Cahn–Hilliard equation with smooth initial-value $u_0 = A^{-1}v \in \dot{H}^2$. To prove this we adapt the arguments in [34, Chapt. 3] and [12, Sect. 5], where error estimates with lower initial regularity are proved. Let $u(t) = E(t)u_0$, $u_h(t) = E_h(t)P_h u_0$. Then the error $e = u_h - u$ satisfies the equation, see [12, (5.4)],

$$(5.6) \quad G_h^2 \dot{e} + e = \rho + G_h \eta, \quad t > 0; \quad P_h e(0) = 0,$$

with

$$(5.7) \quad G_h = A_h^{-1}P_h, \quad \rho = (R_h - I)u, \quad \eta = -(R_h - I)A^{-1}\dot{u}.$$

Any solution of an equation of the form (5.6) satisfies the following bound, with arbitrary $\epsilon > 0$,

$$(5.8) \quad \|e(t)\| \leq \epsilon \sup_{s \in [0, t]} s \|\dot{\rho}(s)\| + C_\epsilon \sup_{s \in [0, t]} \|\rho(s)\| + \left(\int_0^t \|\eta(s)\|^2 ds \right)^{1/2}.$$

In order to prove this we let e_1 be the solution of (5.6) with only ρ as the source term and $P_h e_1(0) = 0$. Moreover, we let e_2 solve the same equation but driven by $G_h \eta$ alone and with $e_2(0) = 0$. Then $e = e_1 + e_2$ solves (5.6). We quote a bound for e_1 from [34, Lemma 3.5]:

$$\|e_1(t)\| \leq \epsilon \sup_{s \in [0, t]} s \|\dot{\rho}(s)\| + C_\epsilon \sup_{s \in [0, t]} \|\rho(s)\|.$$

In order to quote this lemma we note that G_h is self-adjoint, positive semidefinite on \dot{H} and that $G_h e_1(0) = A_h^{-1}P_h e_1(0) = 0$. For e_2 we have

$$\|e_2(t)\| \leq \left(\int_0^t \|\eta(s)\|^2 ds \right)^{1/2}.$$

This is proved by a simple energy argument, see the beginning of the proof of [12, Lemma 5.2]. The reason why we need different proofs for e_1 and e_2 is that G_h in front of η must not appear in (5.8) for we have good bounds for η but not for $G_h \eta$.

This proves (5.8), which can now be combined with bounds for ρ and η , obtained from bounds for R_h and regularity estimates for $u = E(t)u_0$, to get an error bound for $\|e(t)\|$. However, we aim for $\|\dot{e}(t)\|$ and therefore take the derivative of the equation in (5.6) and multiply by t to obtain an equation for $t\dot{e}(t)$:

$$tG_h^2\ddot{e} + t\dot{e} = t\dot{\rho} + tG_h\dot{\eta},$$

which can be written as

$$G_h^2(t\dot{e}) + (t\dot{e}) = G_h^2\dot{e} + t\dot{\rho} + G_h(t\dot{\eta}) = -e + \rho + t\dot{\rho} + G_h(\eta + t\dot{\eta}),$$

where we substituted $G_h^2\dot{e} = -e + \rho + G_h\eta$ from (5.6). Thus, $t\dot{e}$ satisfies an equation of the form (5.6) but with ρ and η replaced by $-e + \rho + t\dot{\rho}$ and $\eta + t\dot{\eta}$. An application of (5.8) with $\epsilon = \frac{1}{2}$, say, gives

$$\begin{aligned} t\|\dot{e}(t)\| &\leq \frac{1}{2} \sup_{s \in [0, t]} (s\|\dot{e}(s)\| + 2s\|\dot{\rho}(s)\| + s^2\|\ddot{\rho}(s)\|) \\ &\quad + C \sup_{s \in [0, t]} (\|e(s)\| + \|\rho(s)\| + s\|\dot{\rho}(s)\|) \\ &\quad + \left(\int_0^t (\|\eta(s)\|^2 + s^2\|\dot{\eta}(s)\|^2) ds \right)^{1/2}. \end{aligned}$$

Since t is arbitrary here we may apply a standard kick-back argument to remove the term $s\|\dot{e}(s)\|$. Another application of (5.8), now with $\epsilon = 1$, takes care of the term $\|e(s)\|$, which leads to

$$\begin{aligned} t\|\dot{e}(t)\| &\leq C \sup_{s \in [0, t]} (\|\rho(s)\| + s\|\dot{\rho}(s)\| + s^2\|\ddot{\rho}(s)\|) \\ &\quad + C \left(\int_0^t (\|\eta(s)\|^2 + s^2\|\dot{\eta}(s)\|^2) ds \right)^{1/2}. \end{aligned}$$

Here we use (2.11) and recall the regularity estimates (2.6), (2.7) for $u(t) = E(t)A^{-1}v$:

$$s^j \|D_t^j \rho(s)\| = s^j \|(R_h - I)D_s^j u(s)\| \leq Ch^2 s^j \|AD_s^j E(s)A^{-1}v\| \leq Ch^2 \|v\|$$

and

$$\begin{aligned} \left(\int_0^t s^{2j} \|D_s^j \eta(s)\|^2 ds \right)^{1/2} &= \left(\int_0^t s^{2j} \|(R_h - I)A^{-1}D_s^j u(s)\|^2 ds \right)^{1/2} \\ &\leq Ch^2 \left(\int_0^t s^{2j} \|D_s^{j+1} u(s)\|^2 ds \right)^{1/2} \\ &\leq Ch^2 \left(\int_0^t s^{2j} \|A^{2j} E(s)A^{-1}v\|^2 ds \right)^{1/2} \\ &\leq Ch^2 \left(\int_0^t s^{2j} \|A^{2j+1} E(s)v\|^2 ds \right)^{1/2} \leq Ch^2 \|v\|. \end{aligned}$$

This completes the proof. \square

Corollary 5.3. *Let $0 < \delta < 1$. Then, the following deterministic error estimates hold for $v \in H$:*

$$(5.9) \quad \|A_h R_{k,h}^n P_h v - A_h E_h(t_n) P_h v\| \leq C k^{1/2(1-\delta)} t_n^{-1+\frac{\delta}{2}} \|v\|,$$

$$(5.10) \quad \|A_h E_h(t) P_h v - A E(t) v\| \leq C h^{2(1-\delta)} t^{-1+\frac{\delta}{2}} \|v\|.$$

Proof. The estimates follow from Lemma 5.2 and the bounds

$$\|A_h R_{k,h}^n P_h\| + \|A_h E_h(t_n) P_h\| \leq C t_n^{-1/2}, \quad \|A_h E_h(t) P_h\| + \|A E(t)\| \leq C t^{-1/2}.$$

\square

Theorem 5.4. *Suppose that (4.11) holds, $\|A^{1/2} Q^{1/2}\|_{\text{HS}} < \infty$ and that $|X_0|_\beta < \infty$ for some $1 \leq \beta \leq 2$. Let $h, k > 0$ small and $0 < \epsilon, \delta < 1$. Then, there is $\Omega_{h,k}^\epsilon \subset \Omega$ with $P(\Omega_{h,k}^\epsilon) > 1 - \epsilon$, $K_T > 0$, and $C = C(T, \epsilon, \delta)$ such that for all $\omega \in \Omega_{h,k}^\epsilon$,*

$$\|X(t_n) - X_h^n\| \leq C \left((h^\beta + k^{\beta/4}) |X_0|_\beta + h^{2(1-\delta)} + k^{1/2(1-\delta)} \right), \quad t_n \in [0, T].$$

Proof. It follows from Theorem 3.1, Proposition 3.2, Theorem 4.3 and Lemma 5.1 that for every $0 < \epsilon, \delta < 1$ and $h, k > 0$ small enough, there is $\Omega_{h,k}^\epsilon \subset \Omega$ with $P(\Omega_{h,k}^\epsilon) > 1 - \epsilon$ and $K_{T,\epsilon,\delta} > 0$ such that

$$(5.11) \quad |X(t)|_1^2 + \|X(t)\|_{L_4}^4 \leq K_{T,\epsilon,\delta}, \quad t \in [0, T], \quad \omega \in \Omega_{h,k}^\epsilon,$$

$$(5.12) \quad \frac{\|X(t) - X(s)\|}{|t - s|^{1/2(1-\delta)}} \leq K_{T,\epsilon,\delta}, \quad t, s \in [0, T], \quad \omega \in \Omega_{h,k}^\epsilon,$$

$$(5.13) \quad |X_h^n|_1^2 + \|X_h^n\|_{L_4}^4 \leq K_{T,\epsilon,\delta}, \quad t_n \in [0, T], \quad \omega \in \Omega_{h,k}^\epsilon$$

$$(5.14) \quad \|W_A(t_n) - W_{A_h}^n\| \leq K_{T,\epsilon,\delta} (h^2 + k^{1/2}), \quad t \in [0, T], \quad \omega \in \Omega_{h,k}^\epsilon.$$

Note that it is enough to establish the above four bounds individually with $\epsilon/4$ on $\Omega_{h,k}^{\epsilon/4,i}$, $i = 1, \dots, 4$, and then set $\Omega_{h,k}^\epsilon = \cap_{i=1}^4 \Omega_{h,k}^{\epsilon/4,i}$. The estimate in (5.12) follows directly from the assumption on the initial data and Proposition 3.2. The rest can be proved using Chebychev's inequality. For example, to prove (5.14), consider

$$F_{h,k} := \frac{\sup_{t_n \in [0, T]} \|W_A(t_n) - W_{A_h}^n\|^2}{h^4 + k}.$$

Then, applying Chebychev's inequality and Lemma 5.1,

$$\mathbb{P}\left(\{\omega \in \Omega : F_{h,k} > \alpha\}\right) \leq \frac{1}{\alpha} \mathbf{E}[F_{h,k}] \leq \frac{K_T}{\alpha}.$$

We choose $\alpha = \epsilon^{-1} K_T$ and set $\Omega_{h,k}^\epsilon = \{\omega \in \Omega : F_{h,k} \leq \epsilon^{-1} K_T\}$. Then

$$\mathbb{P}(\Omega_{h,k}^\epsilon) = 1 - \mathbb{P}\left(\{\omega \in \Omega : F_h > \alpha\}\right) \geq 1 - \epsilon,$$

and the assertion follows. Let $\omega \in \Omega_{h,k}^\epsilon$. We decompose the error function e_n as

$$\begin{aligned} e_n &:= X(t_n) - X_h^n = \left(E(t_n) - R_{k,h}^n\right) P_h X_0 \\ (5.15) \quad &+ \sum_{j=1}^n \int_{t_{j-1}}^{t_j} A_h R_{k,h}^{n-j+1} f(X_h^j) - AE(t_n - s) f(X(s)) \, ds \\ &+ W_A(t_n) - W_{A_h}^n =: e_n^1 + e_n^2 + e_n^3. \end{aligned}$$

For the first error term we use (5.3) to get

$$\|e_n^1\| \leq C(h^\beta + k^{\beta/4}) |X_0|_\beta.$$

The term e_n^2 is most involved and we decompose it further as

$$\begin{aligned} &A_h R_{k,h}^{n-j+1} f(X_h^j) - AE_h(t_n - s) f(X(s)) \\ &= \left((A_h R_{k,h}^{n-j+1} - A_h E_h(t_n - t_{j-1})) f(X_h^j)\right) \\ &+ \left((A_h E_h(t_n - t_{j-1}) - AE(t_n - t_{j-1})) f(X_h^j)\right) \\ (5.16) \quad &+ \left(AE(t_n - t_{j-1}) (f(X_h^j) - f(X(t_j)))\right) \\ &+ \left(A(E(t_n - t_{j-1}) - E(t_n - s)) f(X(t_j))\right) \\ &+ \left(AE(t_n - s) (f(X(t_j)) - f(X(s)))\right) \\ &=: e_{n,j}^{2,1} + e_{n,j}^{2,2} + e_{n,j}^{2,3} + e_{n,j}^{2,4} + e_{n,j}^{2,5}. \end{aligned}$$

Further, by Sobolev's and Poincaré's inequality (keeping in mind that $X_h^j \in \dot{H}$), it follows from (5.13) that

$$\begin{aligned} \|e_{n,j}^{2,1}\| &\leq C k^{1/2(1-\delta)} t_{n-j+1}^{-1+\frac{\delta}{2}} \|f(X_h^j)\| \\ &\leq C k^{1/2(1-\delta)} t_{n-j+1}^{-1+\frac{\delta}{2}} (1 + \|X_h^j\|_{L_6}^3) \\ &\leq C k^{1/2(1-\delta)} t_{n-j+1}^{-1+\frac{\delta}{2}} (1 + |X_h^j|_1^3) \\ &\leq C(1 + (K_{T,\epsilon,\delta})^{3/2}) k^{1/2(1-\delta)} t_{n-j+1}^{-1+\frac{\delta}{2}}. \end{aligned}$$

Similarly,

$$(5.17) \quad \|e_{n,j}^{2,2}\| \leq C(1 + (K_{T,\epsilon,\delta})^{3/2}) h^{2(1-\delta)} t_{n-j+1}^{-1+\frac{\delta}{2}}$$

Also, using (5.11) and (5.13),

$$\begin{aligned} (5.18) \quad \|e_{n,j}^{2,3}\| &\leq C(t_n - t_{j-1})^{-3/4} |P(f(X_h^j) - f(X(t_j)))|_{-1} \\ &\leq C t_{n-j+1}^{-3/4} (1 + |X_h^j|_1^2 + |X(t_j)|_1^2) \|X_h^j - X(t_j)\| \\ &\leq C(1 + K_{T,\epsilon,\delta}) t_{n-j+1}^{-3/4} \|X_h^j - X(t_j)\| = C(1 + K_{T,\epsilon,\delta}) t_{n-j+1}^{-3/4} \|e_j\|. \end{aligned}$$

Furthermore, for $s \in [t_{j-1}, t_j]$,

$$\begin{aligned}
 \|e_{n,j}^{2,4}\| &\leq C(s - t_{j-1})^{1/2(1-\delta)}(t_n - s)^{-1+\frac{\delta}{2}}\|f(X(t_j))\| \\
 (5.19) \quad &\leq C(1 + (K_{T,\epsilon,\delta})^{3/2})(s - t_{j-1})^{1/2(1-\delta)}(t_n - s)^{-1+\frac{\delta}{2}} \\
 &\leq C(1 + (K_{T,\epsilon,\delta})^{3/2})k^{1/2(1-\delta)}(t_n - s)^{-1+\frac{\delta}{2}}
 \end{aligned}$$

Finally, using also (5.12), for $s \in [t_{j-1}, t_j]$,

$$\begin{aligned}
 \|e_{n,j}^{2,5}\| &\leq C(t_n - s)^{-3/4}(1 + |X(t_j)|_1^2 + |X(s)|_1^2)\|X(t_j) - X(s)\| \\
 (5.20) \quad &\leq CK_{T,\epsilon,\delta}(1 + K_{T,\epsilon,\delta})(t_n - s)^{-3/4}(t_j - s)^{1/2(1-\delta)} \\
 &\leq CK_{T,\epsilon,\delta}(1 + K_{T,\epsilon,\delta})(t_n - s)^{-3/4}k^{1/2(1-\delta)}.
 \end{aligned}$$

As, by (5.14),

$$\|e_n^3\| \leq K_{T,\epsilon,\delta}(h^2 + k^{1/2}),$$

collecting all the above terms and applying a generalized version of Gronwall's lemma, [12, Lemma 7.1], finishes the proof if k is small enough. \square

Theorem 5.5. *Under the hypothesis of Theorem 5.4 with $\beta = 1$,*

$$\lim_{h,k \rightarrow 0} \mathbf{E} \sup_{t_n \in [0,T]} \|X(t_n) - X_h^n\|^2 = 0.$$

Proof. It follows from Theorem 3.1 and Theorem 4.3 that there is $L_T > 0$ such that

$$\mathbf{E} \sup_{0 \leq t_n \leq T} (\|X(t_n)\|_{L_4}^4 + \|X_h^n\|_{L_4}^4) < L_T.$$

Let $\epsilon > 0$, $0 < h, k < 1$ small enough, and let K_T and $\Omega_{h,k}^\epsilon$ as in Theorem 5.4. Then, using Theorem 5.4 with $\beta = 1$ and $\delta = \frac{1}{2}$, we get

$$\begin{aligned}
 \mathbf{E} \sup_{0 \leq t_n \leq T} \|X(t_n) - X_h^n\|^2 &\leq \int_{\Omega_{h,k}^\epsilon} \sup_{0 \leq t_n \leq T} \|X(t_n) - X_h^n\|^2 d\mathbb{P} \\
 &\quad + 2 \int_{(\Omega_{h,k}^\epsilon)^c} \sup_{0 \leq t_n \leq T} (\|X(t_n)\|^2 + \|X_h^n\|^2) d\mathbb{P} \\
 &\leq C_\epsilon(h^2 + k^{1/2}) + 4\epsilon^{1/2} \left(\int_{(\Omega_{h,k}^\epsilon)^c} \sup_{0 \leq t_n \leq T} (\|X(t_n)\|^4 + \|X_h^n\|^4) d\mathbb{P} \right)^{1/2} \\
 &\leq C_\epsilon(h^2 + k^{1/2}) + 4\epsilon^{1/2} \left(\mathbf{E} \sup_{0 \leq t_n \leq T} (\|X(t_n)\|^4 + \|X_h^n\|^4) \right)^{1/2} \\
 &\leq C_\epsilon(h^2 + k^{1/2}) + 4\epsilon^{1/2} |\mathcal{D}|^{1/2} \left(\mathbf{E} \sup_{0 \leq t_n \leq T} (\|X(t_n)\|_{L_4}^4 + \|X_h^n\|_{L_4}^4) \right)^{1/2} \\
 &\leq C_\epsilon(h^2 + k^{1/2}) + 4\epsilon^{1/2} |\mathcal{D}|^{1/2} K_T.
 \end{aligned}$$

Let $\eta > 0$. Choose $0 < \epsilon < 1$ such that $8\epsilon^{1/2} |\mathcal{D}|^{1/2} K_T < \frac{\eta}{2}$. Therefore, if $\max(h, k) < \left(\frac{\eta}{4C_\epsilon}\right)^2$, then

$$\mathbf{E} \sup_{0 \leq t_n \leq T} \|X(t_n) - X_h^n\|^2 < \eta,$$

and the proof is complete. \square

REFERENCES

- [1] S. Becker and A. Jentzen, *Strong convergence rates for nonlinearity-truncated Euler-type approximations of stochastic Ginzburg-Landau equations*, ArXiv e-prints (2016).
- [2] D. Blömker and A. Jentzen, *Galerkin approximations for the stochastic Burgers equation*, SIAM J. Numer. Anal. **51** (2013), 694–715.
- [3] D. Blömker, M. Kamrani, and S. M. Hosseini, *Full discretization of the stochastic Burgers equation with correlated noise*, IMA J. Numer. Anal. **33** (2013), 825–848.
- [4] D. Blömker, S. Maier-Paape, and T. Wanner, *Spinodal decomposition for the stochastic Cahn-Hilliard equation*, International Conference on Differential Equations, Vol. 1, 2 (Berlin, 1999), World Sci. Publ., River Edge, NJ, 2000, pp. 1265–1267.
- [5] Z. Brzeźniak, E. Carelli, and A. Prohl, *Finite-element-based discretizations of the incompressible Navier-Stokes equations with multiplicative random forcing*, IMA J. Numer. Anal. **33** (2013), 771–824.
- [6] Caroline Cardon-Weber, *Implicit approximation scheme for the cahn-hilliard stochastic equation*, Preprint, Laboratoire des Probabilités et Modèles Aléatoires, Université Paris V, 2000.
- [7] E. Carelli and A. Prohl, *Rates of convergence for discretizations of the stochastic incompressible Navier-Stokes equations*, SIAM J. Numer. Anal. **50** (2012), 2467–2496.
- [8] H. E. Cook, *Brownian motion in spinodal decomposition*, Acta Metallurgica **18** (1970).
- [9] S. Cox and J. van Neerven, *Pathwise Hölder convergence of the implicit-linear Euler scheme for semi-linear SPDEs with multiplicative noise*, Numer. Math. **125** (2013), 259–345.
- [10] G. Da Prato and A. Debussche, *Stochastic Cahn-Hilliard equation*, Nonlinear Anal. **26** (1996), no. 2, 241–263.
- [11] G. Da Prato and J. Zabczyk, *Stochastic Equations in Infinite Dimensions*, Encyclopedia of Mathematics and its Applications, vol. 44, Cambridge University Press, Cambridge, 1992.
- [12] C. M. Elliott and S. Larsson, *Error estimates with smooth and nonsmooth data for a finite element method for the Cahn-Hilliard equation*, Math. Comp. **58** (1992), 603–630, S33–S36.
- [13] G. B. Folland, *Real Analysis*, second ed., Pure and Applied Mathematics (New York), John Wiley & Sons, Inc., New York, 1999, Modern techniques and their applications, A Wiley-Interscience Publication.
- [14] I. Gyöngy and N. Krylov, *On the rate of convergence of splitting-up approximations for SPDEs*, Stochastic inequalities and applications, Progr. Probab., vol. 56, Birkhäuser, Basel, 2003, pp. 301–321.
- [15] ———, *On the splitting-up method and stochastic partial differential equations*, Ann. Probab. **31** (2003), no. 2, 564–591.
- [16] I. Gyöngy and A. Millet, *On discretization schemes for stochastic evolution equations*, Potential Anal. **23** (2005), 99–134.
- [17] ———, *Rate of convergence of implicit approximations for stochastic evolution equations*, Stochastic Differential Equations: Theory and Applications, Interdiscip. Math. Sci., vol. 2, World Sci. Publ., Hackensack, NJ, 2007, pp. 281–310.
- [18] I. Gyöngy, S. Sabanis, and D. Šiška, *Convergence of tamed Euler schemes for a class of stochastic evolution equations*, Stoch. Partial Differ. Equ. Anal. Comput. **4** (2016), 225–245.
- [19] E. Hausenblas, *Numerical analysis of semilinear stochastic evolution equations in Banach spaces*, J. Comput. Appl. Math. **147** (2002), 485–516.
- [20] ———, *Approximation for semilinear stochastic evolution equations*, Potential Anal. **18** (2003), 141–186.
- [21] M. Hutzenthaler and A. Jentzen, *On a perturbation theory and on strong convergence rates for stochastic ordinary and partial differential equations with non-globally monotone coefficients*, ArXiv e-prints (2014).
- [22] M. Hutzenthaler, A. Jentzen, and D. Salimova, *Strong convergence of full-discrete nonlinearity-truncated accelerated exponential Euler-type approximations for stochastic Kuramoto-Sivashinsky equations*, ArXiv e-prints (2016).
- [23] A. Jentzen and P. Pušnik, *Strong convergence rates for an explicit numerical approximation method for stochastic evolution equations with non-globally Lipschitz continuous nonlinearities*, ArXiv e-prints (2015).
- [24] G. T. Kossioris and G. E. Zouraris, *Finite element approximations for a linear fourth-order parabolic SPDE in two and three space dimensions with additive space-time white noise*, Appl. Numer. Math. **67** (2013), 243–261.

- [25] M. Kovács, S. Larsson, and F. Lindgren, *On the backward Euler approximation of the stochastic Allen-Cahn equation.*, J. Appl. Probab. **52** (2015), 323–338.
- [26] M. Kovács, S. Larsson, and F. Lindgren, *On the discretisation in time of the stochastic Allen-Cahn equation*, ArXiv e-prints (2015).
- [27] M. Kovács, S. Larsson, and A. Mesforush, *Finite element approximation of the Cahn-Hilliard-Cook equation*, SIAM J. Numer. Anal. **49** (2011), 2407–2429.
- [28] ———, *Erratum: Finite element approximation of the Cahn-Hilliard-Cook equation*, SIAM J. Numer. Anal. **52** (2014), 2594–2597.
- [29] Raphael Kruse, *Optimal error estimates of Galerkin finite element methods for stochastic partial differential equations with multiplicative noise*, IMA J. Numer. Anal. **34** (2014), no. 1, 217–251.
- [30] ———, *Strong and weak approximation of semilinear stochastic evolution equations*, Lecture Notes in Mathematics, vol. 2093, Springer, Cham, 2014.
- [31] R. Kurniawan, *Numerical approximations of stochastic partial differential equations with non-globally Lipschitz continuous nonlinearities*, Master thesis, ETH Zürich, 2014, <http://www.sam.math.ethz.ch/teaching/termproj/MasterThesis.RyanKurniawan.pdf>.
- [32] S. Larsson and A. Mesforush, *Finite-element approximation of the linearized Cahn-Hilliard-Cook equation*, IMA J. Numer. Anal. **31** (2011), 1315–1333.
- [33] J. Printems, *On the discretization in time of parabolic stochastic partial differential equations*, M2AN Math. Model. Numer. Anal. **35** (2001), 1055–1078.
- [34] V. Thomée, *Galerkin finite element methods for parabolic problems*, second ed., Springer Series in Computational Mathematics, vol. 25, Springer-Verlag, Berlin, 2006.

CYBERMEDIA CENTER, OSAKA UNIVERSITY, 1-32 MACHIKANNEYAMA, TOYONAKA, OSAKA 560-0043, JAPAN

E-mail address: `furihata@cmc.osaka-u.ac.jp`

MATHEMATICAL SCIENCES, CHALMERS UNIVERSITY OF TECHNOLOGY AND UNIVERSITY OF GOTHENBURG, SE-412 96 GOTHENBURG, SWEDEN

E-mail address: `mihaly@chalmers.se`

MATHEMATICAL SCIENCES, CHALMERS UNIVERSITY OF TECHNOLOGY AND UNIVERSITY OF GOTHENBURG, SE-412 96 GOTHENBURG, SWEDEN

E-mail address: `stig@chalmers.se`

CYBERMEDIA CENTER, OSAKA UNIVERSITY, 1-32 MACHIKANNEYAMA, TOYONAKA, OSAKA 560-0043, JAPAN

E-mail address: `fredrik.lindgren1979@gmail.com`